A Holistic Approach to Integrating Big Data and Strategy: Lessons from the Financial Services Industry

Troy Hamilton
October 10, 2013
Welcome

• SDM 97 Fellow
• Today
• October 25, 2001 – Leadership in a Complex and Changing Business Environment
  • MIT Sloan Professor Hax “How organization changes from “transaction-centric” to “customer-centric” will be key to success”
  • Datek Co-founder & CTO Peter Stern “We had a lot of good ideas, made a lot of mistakes, then tried to understand why. We were then able to change the thinking that led to the mistakes”
• What do our client’s want/need?
Progress is inevitable
we are at a tipping point

SECURITY

PROPRIETARY TECHNOLOGY

NO SINGLE TRUSTED VENDOR

COST PRESSURES

TECHNOLOGY COMMODIZATION

REGULATION MANAGEMENT

NEW CLOUD SERVICES
Duplication adds unnecessary costs
$5,000,000,000,000
A need for a new way of working
simple is better
Client challenges

Client Challenge

- Innovation Constrained
- Staying Ahead of Change
- Managing Data Deluge
- Penetrating New Markets
- Driving Costs Down

Solution

- Open Platform for Collaboration
- Optimized Infrastructure
- Innovative Data Management
- Global Access to Liquidity
- Value Driven Solutions
Enabling Financial Communities Across The World
Our business by the numbers

We execute 8,000,000 trades in all asset classes every day

1.4 billion messages
produced daily (3.5x more than daily Google searches)

One third of all global cash trading executed on NYSE Euronext platforms

Exclusive access for derivatives traders to more than 70 developed, emerging and frontier markets

We run the world’s largest stock exchanges and our exchange technology power more than 16 third-party venues

2,500,000,000,000
financial transactions executed globally every day
(42x more than eBay completes in a year)
A History of anticipating change

1792  Origins of NYSE
1801  Exchange Formalization
1867  The First Stock Ticker
1993  Amex introduces ETFs
1997  Arca ECN Launches

2005  NYSE Hybrid Market
2006-08 Pivotal Mergers
2009  NYSE Technologies
2010  Data Center Launches
2011  NYSE Technologies Capital Markets Community Cloud Launched
2012  NYSE Technologies Introduces OpenMAMA via Linux
Ecosystem

- 2 Data Centers
- 9 Liquidity Centers
- 37 Access Centers & Access Points
I. Open Platform for Collaboration

Challenge:

• Innovation Constrained
  – Security, resiliency, and reliability affect content delivery
  – Highly fragmented community with no central facility
  – Closed platforms are not sustainable
Open platform solutions

**OpenMAMA**
- High-performance, open source Middleware Agnostic Messaging API
- Consistent abstraction layer over message oriented middleware systems
- Removes friction from integrating technologies

**Open Market Data Model**
- A protocol to standardize market data across the industry
- Reduces complexity, risks, and costs of market data management

**Feed Handler Software Development Kit**
- OpenMAMA-compliant framework for integrating data feeds under a single API.
Open platform Initiative

Open Platform Expansion

• Industry standard, data, messaging and application delivery platform through an open source API and SDK
• Enables content crowd-sourcing
• Eliminates vendor lock-in
• Reduces TCO and time to market
Open platform case study

Vendor API

Vendor Middleware

Vendor Normalization

Client Abstraction

Client Applications

MAMA

Data Fabric Middleware

Vendor Normalization

Client Abstraction

Client Applications

In-House API

In-House Middleware

In-House Normalization

Client Application

OpenMAMA – Open Source Standardized Interface

Vendor Middleware

Data Fabric Middleware

In-House Middleware

Open MDM – Open Standardize Normalization Model

Client Applications
Growing participation in openmama
II. Optimized infrastructure

Challenge: Staying Ahead of Change

- Time spent managing legacy infrastructure limits innovation
- Decline in profits forces aggressive cost cutting
- Responding to regulation and security issues
- Current efforts may be commoditized and duplicated enterprise-wide

Clients Looking to Solve Problems with “As A Service” Model
Infrastructure solutions

**SFTI**
- High performance, secure, low latency global network with 1,600 clients and markets
- Powers every NYSE market globally
- Delivers order routing, market data, and access to global hosting locations

**Compute Services**
- Community cloud with compute and storage resources in a managed, secure network environment

**Managed Market Data and Infrastructure Services**
- Ticker Plant deployments in our facilities or client-site
- Hosted and manage servers, switch, firewalls and storage

**Market & Data Center Connectivity**
- Market Connectivity: Access to the network backbone via direct and managed ports
- Data Center Connectivity: Low latency access to NYSE data centers (SFTI IP, SFTI Optic; Wave/Microwave)
Infrastructure initiatives

Commercial Hosting
- Secure, high availability data centers suited to mission critical systems
- Lower cost alternative to proprietary data center build-outs

Asian Expansion
- Managed services, ordering routing and market data for Asian markets
- NYSE provides regional support and local experts

Hosted Networks
- Client access network hosted on SFTI
- Reduces TCO
Infrastructure Case studies

Problem:

- Russell’s approach to index calculation and distribution utilized third parties
- This made them unable to create and market new indices quickly and maintain flexibility

Solution:

- Hosting Russell’s index engine in Mahwah in a fully managed environment
- Market Data via SuperFeed assists Russell in creating indices
- Global Distribution of Russell Tick leveraging our distribution technology and SFTI network

Problem:

- Getco needed to test its algorithms with raw Tokyo market data
- The customer’s international data was too latent by the time it arrived in New Jersey

Solution:

- SFTI provisioned Tokyo market data to Knight’s New Jersey data center quickly
- Latency was still too high, so they began consuming market data in their Tokyo data center via SFTI
III. Innovative Data Management

Challenge:

- **Managing the Data Deluge**
  - Content delivery has not changed and alternatives are needed
  - Cost, ownership, and usage questions remain
$25B SPEND ON MARKET DATA IS PRIME COST OUT TARGET

11.4Bn

- Thomson Reuters
- Bloomberg
- FactSet
- IDC
- Other

9.8Bn

- Thomson Reuters
- Bloomberg
- Other
- IDC
- FactSet

4.3Bn

- Thomson Reuters
- Bloomberg
- Other

(Market data only, including terminals)
Data management solutions

**SuperFeed**
- Consolidated market data feed providing normalized content from 104+ global markets
- Conflated version available
- High Density version includes a fully managed appliance deployed on client site

**Real-time and Historical Data**
- Low-latency feeds ranging from CTA to other data vendors to the cost saving BQT feed
- Historical data includes tick-by-tick data for trades, quotes, and depth of book data

**Direct Feed Handlers & Data Platform Technology**
- 215 low latency, high performance feed handlers for the world’s leading markets
- Data Platform offers scalable, message-oriented middleware for data distribution

**Big Data Software**
- Datatricity takes vast amounts of data from a source, manipulates it, and delivers it for analysis
Data management initiatives

**Symbology & Field Mapping via SuperFeed**
- Aggregated data feed solution offering technology to lower switching costs
- Handles symbology data field transformation via 3rd parties such as Magtia
- Exchange data integration + OTC data

**Tick as a Service**
- Shared utility to capture and stores data across asset class and regions
- Ability to service multiple dedicated client solutions
Data Management Case Studies

Problem:
• EU: Technology refresh of core tickerplant
• US: Reduce costs associated with deployed tickerplant

Solution:
• We designed, built, deployed and managed European tickerplant
• Converted US Tickerplant to SuperFeed to reduce internal charge backs for servers and hosting

Problem:
• Customer encumbered by managing and maintaining in-house and vendor market data systems

Solution:
• We designed, built, deployed and managed all US Equities market data infrastructure.
• Sources European and Futures data from SuperFeed.
• Fully managed service delivery to client via SFTI.
IV. Global access to Liquidity

Challenge: Penetrating New Markets

- High cost of liquidity discovery in fragmented market
- Revenue pressure due to increasing regulatory costs
- Clients need a single point of access to liquidity
- Need for managed solutions that provide normalized connectivity

Source: Electronic Order Book Value (Jan 1 – Feb 28 2011/2010), World Federation of Exchanges
liquidity solutions

**FIX Marketplace**
- Global community of trading counterparties and innovative services
- 750+ buy-sides, 700+ sell-sides
- Services include Order Routing, Liquidity Discovery, and more

**FIX Software**
- Appia software includes a FIX engine, monitoring, and configuration tools
- Market Access Gateways (MAGs) deliver low latency, high throughput access to liquidity venues

**Market Solutions**
- End-to-end order matching solutions
- All client connectivity gateways, on-boarding, UTP matching engine, data distribution, regulatory reporting, and surveillance
liquidity initiatives

**Trade Data Services for Regulatory Reporting**
- Real-time aggregation & normalization of execution drop-copies and storage for regulatory reporting, archive and surveillance.

**OTC Market Access**
- Hosted managed market access for OTC marketplace
- Aggregated access to disparate OTC markets
- Eliminates overhead of EDCs

**FIX as a Service**
- Fully outsourced solution for the sell-side’s entire FIX infrastructure
- Technology investments no longer required and technical staff can be redeployed
Liquidity Case studies

**Problem:**
- State Street ran an expensive, large, and resource-intensive client connectivity network, but it was not a core business nor differentiator

**Solution:**
- Outsourced their inbound client connectivity to us, lowering operating costs, refocusing IT, and accessing new trading partners

**Problem:**
- Deutsche Bank needed a FIX engine to connect to their clients that could scale across regions and asset classes

**Solution:**
- Deutsche Bank now has 80 instances of the Appia FIX Engine implemented with 3000 connections managed globally
V. Value-driven solutions

Challenge:

Drive Down Costs

- Billions of dollars in cost reductions required from infrastructure
- Growing costs around uncontrollable issues, like regulation
- Non-revenue opportunities costing too much time and money

Firms Expected to Cut 20-30% of Costs
Value-driven solutions

**Managed Services**
- 1st Financial Services Cloud
- Access platform services on low latency infrastructure
- Reduce costs and improve time to market

**Data Administration and Reporting Technologies (DART)**
- Manage high volumes of market data and monitor usage
- Reduce overpayment for unused services

**Best Quote and Trades (BQT) Feed**
- Real-time, unified view of BBO, Last Sale data for NYSE Group
- Save up to 45% annually on Level 1 data
Value-driven solutions

Partner as a Service

• Partnering with leading financial software providers to create “As A Service” offerings to further strengthen our product portfolio

• An example is algorithmic back-testing. Our sandbox will allow proprietary data to be integrated with historical market data to validate trade and execution strategies
Value-driven solutions case study

**Problem:**
- Pershing needed to minimize in-house FIX management and reduce connectivity overhead

**Solution:**
- Aggregated buy-side connectivity into an internal OMS - reducing FIX management footprint.
- Customized API into NYXT from an internal OMS to access US markets – eliminating need to keep up with EDCs.
- Direct client access into Pershing’s FX pricing engine
- Access for London clients to EU BATS and back to the US

**Problem:**
- TD Ameritrade needed an agile and flexible Disaster Recovery solution while containing costs.

**Solution:**
- NYSE crafted a virtual data center extension so that the customer only paid for what they used
- NYSE maintains, and operates 200 online broker virtual servers on 45 physical servers in Mahwah
- NYSE provides redundant dark fiber to TD’s primary site SFTI ports
##### NYSE US Markets: Market Data / Surveillance Challenges

**Before Datatricity 2007**

- Massive Daily Volume
- Historical Retention & Data Re-Use
- Expensive Analytics Platforms
- System Complexity
- Slow System Performance
- Analytics based on correlation of disparate data sources
- Heavy Reliance on IT
NYSE US markets:
Scalable efficient Solution

Realized Business Benefits

✓ Self-Service
✓ Scalable data growth management
✓ IT Cost Reduction for Data Operations and Access
✓ Improved Productivity Business Response Time, Technology Agility
✓ Reduced Licensing Costs for Database Infrastructure
✓ 360° View of Users, Resources, Entitlements, Data Mappings, Usage Metrics and Job Status
✓ Centralized Data Access Controls
✓ Simplified access and use of “Big Data”
NYSE Euronext Big Data Challenges

- Globally distributed data using different technologies & formats
- Slow IT response time to Business Analytical needs
- Multiple Analytics Platforms required for different business use cases
- No centralized entitlements for data access
- Need to support rapidly growing data footprint with 5 TB+ daily
- Need for near-real time Analytics
- Rapidly changing market structure translated into frequent data schema changes
- Regulated environment with zero tolerance for data loss
- Limited visibility to data consumer activity and demand
- Multi-vendor, multi-platform environment
Brief History of On-Demand Data

• The On-Demand Data application sprouted late 2007 to solve **NYSE Market Surveillance challenges**. NYSE needed a **more cost-efficient and scalable solution**, one that could be easily managed, to ingest, archive, and access Trading and Market Data. The primary constituents were Compliance, Research and Market Surveillance groups.

• The New York Stock Exchange faced several big data challenges. These included **massive daily volume ingestion, historical data retention and online access requirements**, data analytics platform cost proliferation, long development lifecycle for analytics, data & system complexity manifesting within silos, lack of data integration, difficulty keeping up with shrinking end-of-day operations windows, and over reliance on IT staff.

• On-Demand Data was developed to provide End-Users access to NYSE File Archives and common data services. These included rapid database provisioning, a tool for distributed data consolidation and integration with high-performance Analytics platforms (Greenplum, Netezza, Teradata, etc).
Realized Benefits of NYSE Big Data Platform

- Greatly reduced TCO of Regulatory Data Platforms
- Centralized Management of Distributed Files at Global Scale
- Self Service Data Provisioning & Sandbox for User Analytics
- Self Service Data Filtering & Transformation
- File/Data Movement Workflow & Orchestration
- Optimized Connectors to Big Data & Analytics Platforms
- Simplify Access & Integration of Data
- Single UI to Heterogeneous Systems ("same look-n-feel")
- High Performance
- Managing Schema Change over time
- Managing Data Retention & Data Purge
- Tracking Data Usage & Resource Utilization
- Technology Platform Agnostic
On-Demand Data Platform Services

User Services
- Admins
- Users
- API Applications

File Farms
- Streams
- Hadoop

RDBMS

Clouds

Archive
- Extract
- Generic Loader
- High-Speed Loader
- Copy (Data Ship)
- Workflow
- Data Pump
- De/Compress
- De/Encrypt

On-Demand Data Services
- Linux Cluster
- File Streaming
- Pipeline Filtering
- Pipeline Transformation
- Scheduled/Purge
- Auto Data Expiration
- Metadata Repository
- Worker Load Balancing
- Resource Throttling
- Work Scheduler
- File Access Locking

User-driven Data Provisioning
- Federated Data Store
- Federated User Entitlements
- Usage Tracking
- Service Monitoring
- File Farm Monitoring
- File Search & Load
- Schema Extracts & Archive

Data Analytics Platforms
- Netezza
- Oracle
- Greenplum
- SQL Server
- Hadoop

Big Data Roles
- On-Demand Data System Administrators
- Data Administrators
- Super Users
- Data Consumers

Leadership, Innovation, Systems Thinking
Final Thoughts

• Progress and change are inevitable
• Simplify your clients' lives
• Holistic view - understanding of your clients' data needs
• Customer-centric, not transaction-centric
• Have a systems view to bring it all together
Open platform for collaboration

- **Challenges:**
  - Security, resiliency, and reliability affect content delivery
  - Highly fragmented community with no central facility
  - Closed platforms are not sustainable

- **Solutions:**
  - OpenMAMA
  - Open Market Data Model
  - Feed Handler SDK

- **Benefits:**
  - Removes friction from integrating different technologies
  - Reduces vendor lock-in
  - Reduces complexity, risks, and costs of market data management
OPTIMIZED INFRASTRUCTURE

Challenge:
• Time spent managing legacy infrastructure limits innovation
• Decline in profits forces aggressive cost cutting
• Responding to regulation and security issues
• Current efforts may be commoditized and duplicated enterprise-wide

Solutions:

Benefits:
• Save time and costs by using a secure, high availability network
• Lower cost alternative to proprietary data center build-outs
• Managed solutions reduce time-intensive tasks and comply with new regulations
INNOVATIVE DATA MANAGEMENT

• **Challenge:**
  • Content delivery has not changed and alternatives are needed
  • Cost, ownership, and usage questions remain

• **Solutions:**
  - SuperFeed
  - Real-time & Historical Data
  - Direct Feed Handlers and Data Platform Technology
  - Big Data Software

• **Benefits:**
  • Lower cost consolidated market data feed
  • Broad range of market data solutions available as a service allowing you to focus on your trading strategy
GLOBAL ACCESS TO LIQUIDITY

Challenge:
• High cost of liquidity discovery in fragmented market
• Revenue pressure due to increasing regulatory costs
• Clients need a single point of access to liquidity
• Need for managed solutions that provide normalized connectivity

Solutions:

1. FIX Marketplace
2. FIX Software
3. Market Solutions

Benefits:
• Reduce cost and complexity by accessing a global connected community
• Managed services for scalable and comprehensive trading solutions
VALUE-DRIVEN SOLUTIONS

Challenge:
• Billions of dollars in cost reductions required from infrastructure
• Growing costs around uncontrollable issues, like regulation
• Non-revenue opportunities costing too much time and money

Solutions:
1. Managed Services
2. DART
3. Best Quotes and Trades (BQT) Feed

Benefits:
• Reduce costs and improve time to market with a managed solution
• Avoid overpayment for unused market data
• Lower cost alternatives for market data feeds